



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2014

The effects of stabilizing and directional selection on phenotypic and genotypic variation in a population of RNA enzymes

Hayden, Eric J ; Bratulic, Sinisa ; Koenig, Iwo ; Ferrada, Evandro ; Wagner, Andreas

Abstract: The distribution of variation in a quantitative trait and its underlying distribution of genotypic diversity can both be shaped by stabilizing and directional selection. Understanding either distribution is important, because it determines a population's response to natural selection. Unfortunately, existing theory makes conflicting predictions about how selection shapes these distributions, and very little pertinent experimental evidence exists. Here we study a simple genetic system, an evolving RNA enzyme (ribozyme) in which a combination of high throughput genotyping and measurement of a biochemical phenotype allow us to address this question. We show that directional selection, compared to stabilizing selection, increases the genotypic diversity of an evolving ribozyme population. In contrast, it leaves the variance in the phenotypic trait unchanged.

DOI: <https://doi.org/10.1007/s00239-013-9604-x>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-88300>

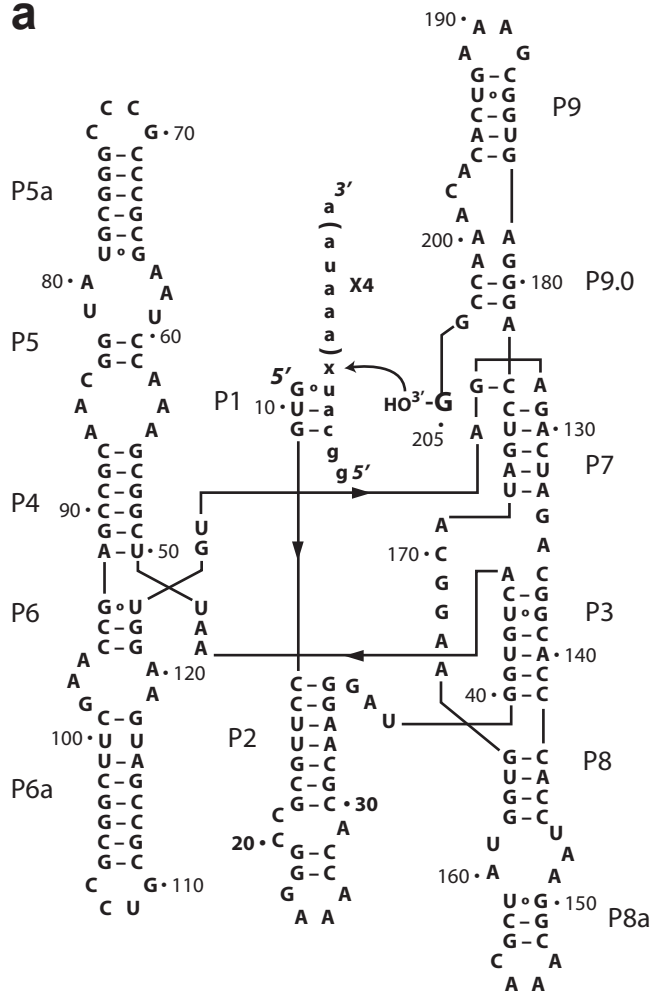
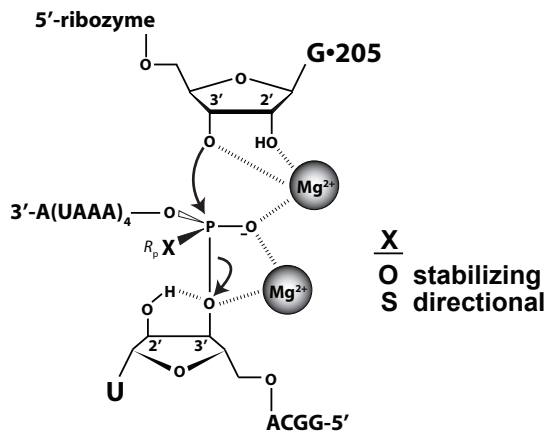
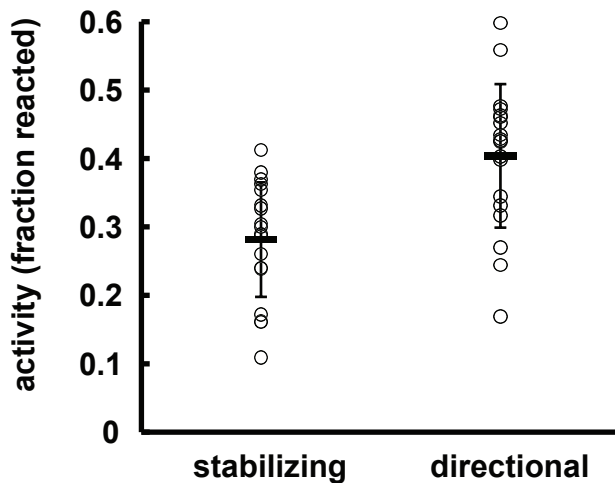
Journal Article

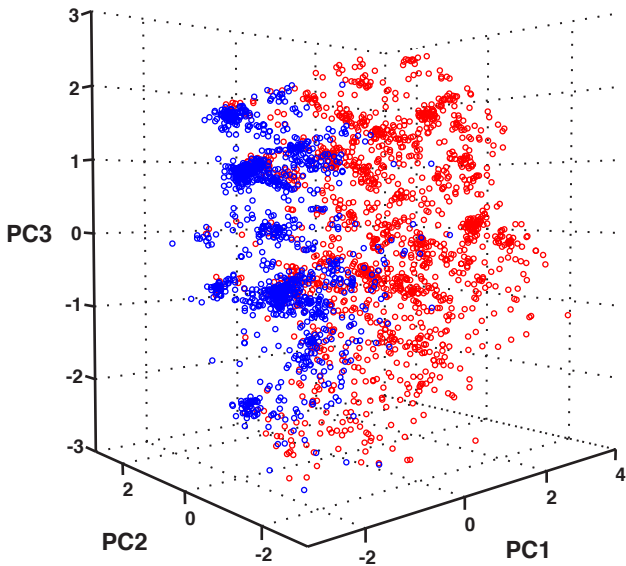
Accepted Version

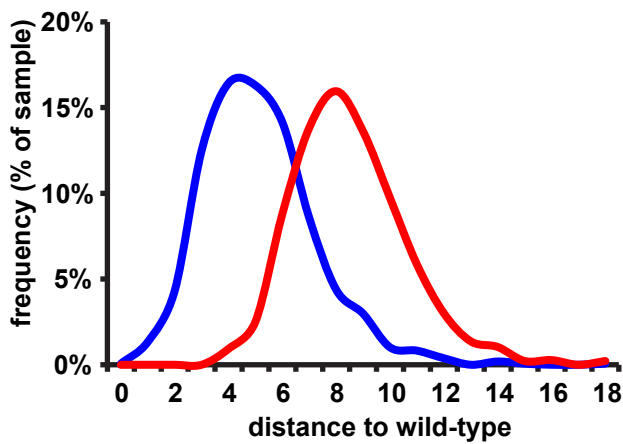
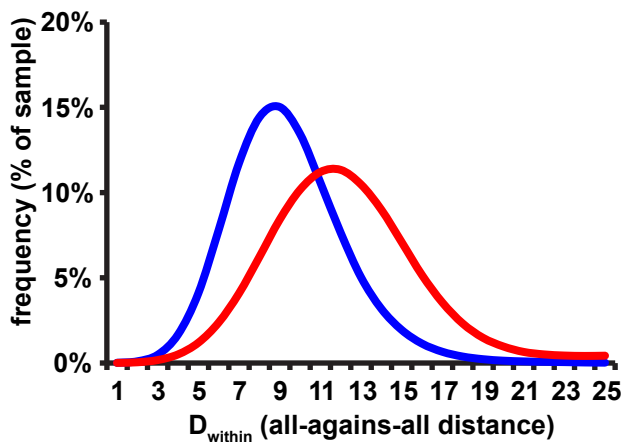
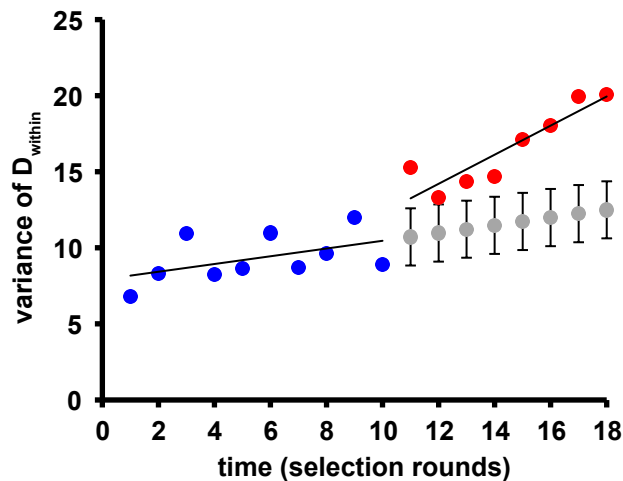
Originally published at:

Hayden, Eric J; Bratulic, Sinisa; Koenig, Iwo; Ferrada, Evandro; Wagner, Andreas (2014). The effects of stabilizing and directional selection on phenotypic and genotypic variation in a population of RNA enzymes. *Journal of Molecular Evolution*, 78(2):101-108.

DOI: <https://doi.org/10.1007/s00239-013-9604-x>

a**b****c**



a**b****c**

The effects of stabilizing and directional selection on phenotypic and genotypic variation in a population of RNA enzymes

Eric J. Hayden¹, Sinisa Bratulic^{2,4}, Iwo Konig³, Evandro Ferrada⁵, and Andreas Wagner^{2,4,5}

¹Department of Biological Science, Boise State University, Boise, Idaho USA, ²Institute of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland, ³Department of Biochemistry, University of Zurich, Zurich, Switzerland, ⁴The Swiss Institute of Bioinformatics, Lausanne, Switzerland, ⁵The Santa Fe Institute, Santa Fe, New Mexico, USA

Corresponding Author:

Eric J. Hayden

email: erichayden@boisestate.edu

Office Number: [208-426-4625](tel:208-426-4625)

Office Fax: [208-426-1040](tel:208-426-1040)

Abstract

The distribution of variation in a quantitative trait and its underlying distribution of genotypic diversity can both be shaped by stabilizing and directional selection. Understanding either distribution is important, because it determines a population's response to natural selection. Unfortunately, existing theory makes conflicting predictions about how selection shapes these distributions, and very little pertinent experimental evidence exists. Here we study a simple genetic system, an evolving RNA enzyme (ribozyme) in which a combination of high throughput genotyping and measurement of a biochemical phenotype allow us to address this question. We show that directional selection, compared to stabilizing selection, increases the genotypic diversity of an evolving ribozyme population. In contrast, it leaves the variance in the phenotypic trait unchanged.

Introduction

A population of organisms or molecules with some quantitative phenotype P , where natural selection favors an increase or decrease in P , and where the population mean of P is consequently changing over time, is said to be under *directional selection*. This mode of selection typically occurs when a population experiences a new environment and many individuals find themselves poorly adapted. Over time, such a population would adapt to its environment, and enter a regime of *stabilizing selection*, during which the population mean phenotype P no longer changes. Because of its importance for understanding Darwinian evolution, the effects of directional and stabilizing selection on phenotypes have received considerable attention by theorists.

One simple question with no straightforward answer regards the effect of directional and stabilizing selection on the phenotypic variance of a population. Because directional selection can favor extreme phenotypes, one might think that it would cause an increase not only in the mean, but also in the phenotypic variance of the population. However, this is not necessarily the case. Whether it does depends on several factors, including the shape – convex or concave – of the fitness function (Layzer 1978), the distribution of trait values (Bürger 1991; Hansen 1992), whether a trait optimum fluctuates with time, the presence and kind of genotype-by-environment interactions (Hermisson and Wagner 2004), the existence of epistatic interactions among loci (Hansen 2006), and the extent of phenotypic canalization before the onset of directional selection (de Visser et al. 2003; Hermisson and Wagner 2004). The change in phenotypic variance caused by periods of stabilizing selection is no easier to predict. It depends upon several attributes of genetic architecture, which include epistasis (Hermisson et al. 2003), the structure of the fitness landscape (Barton and Turelli 1987), the number of

loci involved in a trait (Turelli and Barton 1990; Bürger 1991), and the extent of canalization (Wagner et al. 1997). Overall, existing theory does not make any simple and general predictions for the change in phenotypic variance over time. Unfortunately, there is also very little experimental data that speaks to this question, with the exception of the observation that stressful environments can cause a release of previously hidden phenotypic variation with a genetic basis (Waddington 1959; Scharloo 1991; Rutherford and Lindquist 1998; True and Lindquist 2000; Gibson and Dworkin 2004; Flatt 2005; Hayden et al. 2011).

Directional and stabilizing selection can not only affect the distribution of phenotypes, but also the distribution of genotypes, and especially their diversity. Next to no experimental data is available about these effects, and mainly for two reasons. First, it has not been feasible until recently to genotype many individuals in an evolving population. Second, many phenotypic traits are complex and have unknown genetic components. Next-generation sequencing technologies are beginning to allow us to overcome the first problem (Kvitek and Sherlock 2011; Ensminger et al. 2012; Traverse et al. 2013; Herron and Doebeli 2013), especially for the relatively simple phenotypic traits of molecules, where the second problem is not an issue. For such molecular traits, one can obtain complete genotypic information of many individuals in an evolving population (Hietpas et al. 2011; Hayden et al. 2011). The purpose of this paper is to study the evolution of both phenotypic and genotypic variation in an especially transparent molecular system, a population of evolving RNA enzymes.

The phenotypic trait in our study system is the catalytic activity of the RNA enzyme (ribozyme) derived from the *Azoarcus* group I self-splicing intron (Tanner and Cech 1996). Ribozymes are nucleic acid sequences with a catalytic function. Thus phenotype and genotype are

linked to a single molecule in this system. Populations of ribozyme genotypes can be enzymatically replicated in the lab. Genotypic and phenotypic diversity is introduced through nucleotide mutations made during replication. Selection can be applied to exploit the differences in catalytic activity caused by these mutations, resulting in Darwinian evolution (Joyce 2004). An important feature of our study system is that the phenotypes are encoded in very small genotypes (<200 bp). For such genotypes, high throughput sequence technology can be used to analyze many genotypes in evolving RNA populations (Pitt and Ferre-D'Amare 2010; Hayden et al. 2011). In other words, they allow us to study empirically how such populations spread through the vast genotype space of all possible RNA sequences. This space harbors 4^l possible RNA sequences of length l . Each individual sequence occupies a discrete point in this space. Two sequences have a *distance* that is defined as the number of nucleotide changes necessary to change one sequence into the other. The genotypes of a given population can be viewed as a cloud in genotype space. Here we describe the effects on this cloud of genotypes as it experiences stabilizing selection that is followed by a period of directional selection.

We previously conducted evolution experiments where we exposed the *Azoarcus* group I ribozyme to two different chemical environments (Hayden et al. 2011). The ribozyme has the ability to cleave an exogenous RNA substrate because this activity is a component of its catalytic role in nature as a self-splicing intron (Tanner and Cech 1996). In the first environment we supplied a ribozyme population with a “native” RNA oligonucleotide. Cleavage of this oligonucleotide results in the ligation of a portion the substrate to the 3'-end of the ribozyme, and we can exploit this sequence modification to selectively amplify catalytically successful molecules (Beaudry and Joyce 1992; Lehman and Joyce 1993). Molecules with higher activity are more likely to be replicated. We

previously observed that in this chemical environment, the average activity of the population did not change significantly over eight rounds of mutation and selection on this native substrate. Further, we did not detect any mutations that were increasing in frequency over time, which would indicate a selective advantage for ribozymes containing them. We concluded that the population was well-adapted to this native substrate. Herein, we refer to the population from the eighth round of this selection as our Stabilizing Selection population (SS).

The second chemical environment is one to which the ancestral wild-type *Azoarcus* ribozyme is poorly adapted. Specifically, we introduced a chemical modification (phosphorothioate) into the RNA backbone of the substrate, and exposed a ribozyme population to this modified substrate. We mutagenized the members of this population, and then exerted selection by amplifying molecules that could cleave this new substrate, performing an additional eight “generations” of mutation and selection in this way. Under this second selection pressure, the population adapted rapidly to the new substrate. The average activity of the population after eight generations was significantly higher than it was prior to this selection. This demonstrates directional selection exerted by the introduction of the new substrate. We will refer to the population from the eighth round of this second selection as the Directional Selection (DS) population. These experimentally derived populations facilitate a comparison between the distribution of phenotypes and genotypes resulting from an equal number of rounds (8) of each type of selection. Our previous activity measurements were based on the average activity of the mixed populations, and therefore did not tell us about the variance in activity between different genotypes. We will next describe the distribution of phenotypic values in the SS and DS population, before analyzing the genotypic variation.

RESULTS

Phenotypic variation

To characterize the distribution of phenotypes we isolated 19 individual molecules from each population through molecular cloning (see Methods). We measured the activity of each cloned sequence under the conditions used for selection in that population, i.e. the *native* RNA substrate for the SS population and the *new* phosphorothioate substrate for the DS population. For our phenotypic trait, which we will refer to as *activity*, we determined the fraction of ribozyme that had reacted after 1h with the appropriate substrate (Figure 1).

With the *native* substrate, the distribution of activity values in the SS population after 8 generations had a mean of 0.28 ± 0.04 (mean \pm 95%C.I.), which is very close to that of the wild-type (mean = 0.27 ± 0.03). This indicates that the population indeed experienced stabilizing selection, because its activity did not increase. In contrast, with the *new* substrate, the activity values of the DS population (mean = 0.40 ± 0.05) were nearly double that of the wild-type (mean = 0.21 ± 0.03). In fact, none of the ribozymes sampled from the DS population had activity significantly lower than the wild-type after 8 generations of directional selection. The difference in mean activity between the DS and the SS population was highly statistically significant (t-test; $P = 0.0003$), confirming that the population responded with a phenotypic change to the new chemical environment. Importantly, the variance in the distribution of activity values was not significantly higher in the DS population than in the SS population (F-test $P = 0.27$; $n=38$), and the coefficients of variation were also similar (SS = 0.30, DS = 0.26). In sum, our data show that exposure to the *new* substrate resulted in directional

selection, but that this directional selection did not cause an increase in the variance of phenotypes in the population.

Genotypic variation

To compare the genotypes resulting from stabilizing and directional selection, we analyzed DNA sequence data of 1577 molecules from the SS population, and 1830 molecules from the DS population. The sequence data was obtained as part of a previous publication, where it was partially analyzed to detect high fitness variants (Hayden et al. 2011). Here we study the regions of multidimensional sequence space occupied by the sequences in our two populations.

As an exploratory tool, we used a principle component analysis (PCA) to visualize our high dimensional data from the most informative low dimensional perspectives. We carried out a PCA of a multiple sequence alignment comprised of all the sequence data from both our populations combined (see Methods). Figure 2 shows the data as a three-dimensional scatter plot projection of our populations in sequence space. The plot indicates that the DS population is more genotypically diverse, because the data from this population covers a larger region in the graph than the SS population. We note that this greater diversity occurs despite a several fold lower mutational supply rate in the DS population ($\mu \approx 0.14$ mutation per individual per generation) than in the SS population ($\mu \approx 1.0$) (Hayden et al. 2011).

We next confirmed the visual suggestion of the PCA through a quantitative analysis of the variance in the mutational distance among genotypes. First, we used the wild-type *Azoarcus* ribozyme sequence as a reference point, and determined, for each sequence in both populations, the mutational

distance to the wild-type ribozyme (Figure 3a). This metric confirms the greater genetic diversity in the DS population. We find that the mean distance to wild-type is higher for the DS population (mean = 8.68) than the SS population (mean 5.16). Also, the distribution of distances is broader in the DS population (stdev = 2.58) than the SS population (stdev = 2.13). Both the skew and kurtosis of the distributions are positive for both populations, but both are higher for the DS population (skew = 1.98, kurt = 7.67) than for the SS population (skew = 1.03, kurt = 3.00).

A different measure of genetic diversity is the mutational distance D_{within} between all sequence pairs of a population (all-against-all comparison). For example, two sequences in a population that are each different from the wild-type sequence at two positions (distance = 2) can differ from each other in anywhere between zero and four positions (distance = 0-4). While the distance from the wild-type gives a measure of diversity relative to one reference point, the all-against-all comparison provides a measure of how far a population has spread through genotype space. The distributions of D_{within} measurements for each population are shown in Figure 3b. We find the mean D_{within} is higher in the DS population (12.4) than in the SS population (9.39). Importantly, the variance is more than two times higher in the DS population (20.1) than in the SS population (9.63). This provides quantitative confirmation that the DS population is spread out over a larger region of multi-dimensional space, as had been visually suggested by the PCA.

The comparison of the variance in D_{within} is complicated by the fact that the two populations are related by descent, and therefore the DS population has had more time to accumulate genetic variation. The observed broader distribution of D_{within} for the DS population could result if the increase in variance among genotypes over time were the same under both stabilizing and directional selection. In

order to compare the rate of change of D_{within} during stabilizing and directional selection, we determined the variance in D_{within} over the entire course of both selection regimes (Figure 3c). The results show that the variance was increasing during both selection regimes, as indicated by the positive slopes of both the blue and red data points. However, linear regression analysis confirms that the two slopes are significantly different ($p = 0.0155$, $t = -2.75$). The rate of increase in variance during directional selection (slope = 0.96) is significantly higher than during stabilizing selection (slope = 0.25). For visual comparison, the grey data points in Fig. 3c show the variance in D_{within} that would be predicted from the trend of the stabilizing selection (blue) data.

DISCUSSION

Our experimental system is very simple, which is an advantage compared to the complex genotypes and phenotypes of whole organisms. Some factors that could potentially influence and confound the observed changes in phenotypes and genotypes are not relevant in such a simple system. For example, we can disregard any confounding influence of fluctuating environments, because during a given selection regime in our experiments, we can keep the laboratory environment (temperature, pH, substrate availability, population size) constant, especially when compared to natural environments. Furthermore, each RNA molecule represents a compact haploid genome. We can therefore disregard any complicating factors due to ploidy, such as dominance or overdominance.

Despite its relative simplicity, however, several aspects of our system mirror features of more complex genetic architectures involving many genes. For example, its activity-based phenotype (fraction of ribozyme reacted) can be affected by several fundamental properties of RNA structure and

catalysis, including RNA folding, thermodynamic stability of the catalytically active structure, forward and reverse rates of catalysis, and affinity of the ribozyme for the substrate and product of the reaction. Because usually not all of these properties can be simultaneously optimized in any one molecule, many mutations in enzymes have pleiotropic effects, just as in more complex systems (Wang et al. 2002; Tokuriki et al. 2008; Soskine and Tawfik 2010). In addition, the structural stability of the *Azoarcus* ribozyme enables considerable mutational robustness which allows cryptic genetic variation with little phenotypic effect to accumulate in populations (Hayden et al. 2011). We have previously shown that this cryptic variation can enable faster adaptive evolution, and that robustness can decrease following selection (Hayden et al. 2012), with certain sequences selected for increased activity under laboratory conditions becoming decanalized relative to the wild-type sequence. We note that our current observation of constant phenotypic variance in the face of increased genetic diversity is nevertheless consistent with our previous report of decanalization of a single high-fitness individual (Hayden et al. 2012). There are two main points to consider. First, our previous decanalization experiments involved the introduction of additional random mutations in order to observe the extent of their average deleterious effect, whereas the current analysis examines only the variation maintained in the population in the presence of selection. Second, we previously found that our study population as a whole had only a slight decrease in robustness, demonstrating that the decanalization observed in the single high-fitness variant was offset by increased robustness in other members of the population.

Another parallel between our system and more complex genetic systems is that adaptive mutations within an individual ribozyme sequence can show epistatic interactions (Hayden and Wagner 2012). Similar epistatic interactions also pervade complex genetic systems (Elena and Lenski

Kommentar [a1]: Flow and connection a bit unclear. Write "...", and for two reasons"?

2001; Kvitek and Sherlock 2011; Chou et al. 2011; Khan et al. 2011). This point in particular may help explain our observation that directional selection resulted in an increase in genotypic differences, but this did not cause a corresponding increase in phenotypic variance. Specifically, it is consistent with theoretical observations that epistatic interactions can lead to decreased additive genetic variance (Hermisson et al. 2003). Theoretical arguments like this, together with our past experimental results (Hayden et al. 2011; Hayden and Wagner 2012; Hayden et al. 2012) suggest that phenomena such as hidden genetic variation, mutational robustness and epistatic interactions can influence the distributions of phenotypes and genotypes in our populations.

We note that many of the factors mentioned above are system-specific, and could differ in other experimental systems, where directional selection may indeed increase phenotypic variance. This is important because the rate and nature of evolutionary change depends upon the presence of natural selection and the variation in the selected phenotypic trait (Lynch and Walsh 1998). Understanding the change in heritable phenotypic variation caused by selection is critical for determining the dynamics of populations responding to fluctuating/heterogeneous environments, and for predicting the fate of populations following environmental change. Because empirical data of simultaneous phenotypic and genotypic distributions is virtually absent, more work on tractable systems like ours is desirable. Fortunately, high throughput genotyping techniques will make such work feasible.

Methods

Ribozyme synthesis

The previously reported evolution protocol produced dsDNA of selected RNA molecules (Hayden et

al. 2011). This dsDNA was cloned into the vector pJET1.2 using the CloneJET cloning kit (Fermentas) and transformed into chemically competent *E. coli* (TOP10, Invitrogen). Individual colonies were screened by colony PCR reactions using primers complimentary to the ribozyme sequence, and that also contained an extension to add the promoter for T7 RNA polymerase (written 5'-3'):

Forward primer = CTGCAGAATTCTAATACGACTCACTATAGTGCCTTGCGCCGGGAA

Reverse primer = CCGGTTTGTGTGACTTTCGCC

PCR products of the correct size were ethanol precipitated and rehydrated in TE buffer (pH8) for in vitro transcription using T7 RNA polymerase (Fermentas) using the manufacturer supplied buffer and each NTP at a concentration of 2 mM. RNA was gel-purified by denaturing PAGE (6% polyacrylamide, 8 M urea), visualized by UV light, excised from the gel, and eluted into 0.3M sodium acetate. The eluent was filtered (MF device, Pall scientific), ethanol precipitated, and rehydrated in nuclease free water (Ambion). A sample of the RNA (2 μ L) was then quantified by UV absorbance (Nanodrop).

Activity measurements

Ribozyme activity was measured in 10 μ L reactions containing 10 pmol ribozyme, 25mM MgCl₂, 30 mM EPPS (pH 7.5), and either 25 pmol RNA substrate (SS population) or 50 pmol phosphorothioate substrate (DS population). Reactions were incubated for 1h at 37°C. Reacted and unreacted RNA molecules were separated on a 6% denaturing polyacrylamide gel, which were then stained with GelRed (Biotum) for imaging. Fraction reacted was determined as the fluorescent signal from the reacted band divided by the total fluorescent signal (unreacted plus reacted). Each data point in Fig. 1 represent the average of three replicates.

Formatiert: Einzug: Erste Zeile:
0.74 cm

Sequencing

Sequence data was obtained for a previous publication (Hayden et al. 2011). Briefly, each population was PCR amplified with primers containing six nucleotides unique for each population and the Roche 454 adaptor sequences. Populations were mixed, loaded onto a single picotiter plate divided into 4 regions, and sequenced by pyrosequencing (Roche 454).

Alignments

Distance measurements were obtained by pairwise alignment of each sequence to either the wild-type sequence (distance to wild-type) or to another sequence from the population (all-against-all comparisons). The first 25 and final 21 nucleotides (5'-3') of each sequence were primer binding sites, and nucleotide differences in these regions were excluded from the analysis. Alignments were carried out using USEARCH (Edgar 2010) and MUSCLE (Edgar 2004).

PCA

The sequence data from both populations, as well as from a wild-type reference sequence, were combined, and a large multiple sequence alignment was generated using the program MUSCLE. Sequences that covered less than 95% of the wild-type sequence were discarded. The alignment was trimmed from both ends to the positions that aligned to the portion of the wild-type sequence residing inside the primer binding sites. Sequences were converted to numerical format by replacing A with 1, G with 2, C with 3, T with 4 and gaps with 0. Principle component analysis was conducted with the *princomp* function in Matlab.

Acknowledgments

EJH acknowledges support through the Biomolecular Science Program at Boise State University. AW acknowledges support through Swiss National Science Foundation grant 315230-129708, as well as through the University Priority Research Program in Evolutionary Biology at the University of Zurich.

References

- Barton NH, Turelli M (1987) Adaptive landscapes, genetic distance and the evolution of quantitative characters. *Genet Res* 49:157–173.
- Beaudry AA, Joyce GF (1992) Directed evolution of an RNA enzyme. *Science* 257:635–641.
- Bürger R (1991) Moments, cumulants, and polygenic dynamics. *J Math Biol* 30:199–213.
- Chou H-H, Chiu H-C, Delaney NF, et al. (2011) Diminishing returns epistasis among beneficial mutations decelerates adaptation. *Science* 332:1190–1192. doi: 10.1126/science.1203799
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797. doi: 10.1093/nar/gkh340
- Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26:2460–2461. doi: 10.1093/bioinformatics/btq461
- Elena SF, Lenski RE (2001) Epistasis between new mutations and genetic background and a test of genetic canalization. *Evolution* 55:1746–1752.
- Ensminger AW, Yassin Y, Miron A, Isberg RR (2012) Experimental Evolution of *Legionella pneumophila* in Mouse Macrophages Leads to Strains with Altered Determinants of Environmental Survival. *PLoS Pathog* 8:e1002731. doi: 10.1371/journal.ppat.1002731
- Flatt T (2005) The evolutionary genetics of canalization. *Q Rev Biol* 80:287–316.
- Gibson G, Dworkin I (2004) Uncovering cryptic genetic variation. *Nat Rev Genet* 5:681–690. doi: 10.1038/nrg1426

- Hansen TF (2006) The Evolution of Genetic Architecture. *Annual Review of Ecology, Evolution, and Systematics* 37:123–157. doi: 10.1146/annurev.ecolsys.37.091305.110224
- Hansen TF (1992) Selection in asexual populations: An extension of the fundamental theorem. *Journal of Theoretical Biology* 155:537–544. doi: 10.1016/S0022-5193(05)80634-4
- Hayden EJ, Ferrada E, Wagner A (2011) Cryptic genetic variation promotes rapid evolutionary adaptation in an RNA enzyme. *Nature* 474:92–95. doi: 10.1038/nature10083
- Hayden EJ, Wagner A (2012) Environmental change exposes beneficial epistatic interactions in a catalytic RNA. *Proc R Soc B* 279:3418–3425. doi: 10.1098/rspb.2012.0956
- Hayden EJ, Weikert C, Wagner A (2012) Directional Selection Causes Decanalization in a Group I Ribozyme. *PLoS ONE* 7:e45351. doi: 10.1371/journal.pone.0045351
- Hermisson J, Hansen TF, Wagner GP (2003) Epistasis in polygenic traits and the evolution of genetic architecture under stabilizing selection. *Am Nat* 161:708–734. doi: 10.1086/374204
- Hermisson J, Wagner GP (2004) The population genetic theory of hidden variation and genetic robustness. *Genetics* 168:2271–2284. doi: 10.1534/genetics.104.029173
- Herron MD, Doebeli M (2013) Parallel Evolutionary Dynamics of Adaptive Diversification in *Escherichia coli*. *PLoS Biol* 11:e1001490. doi: 10.1371/journal.pbio.1001490
- Hietpas RT, Jensen JD, Bolon DNA (2011) Experimental illumination of a fitness landscape. *PNAS* 108:7896–7901. doi: 10.1073/pnas.1016024108
- Joyce GF (2004) Directed evolution of nucleic acid enzymes. *Annu Rev Biochem* 73:791–836. doi: 10.1146/annurev.biochem.73.011303.073717
- Khan AI, Dinh DM, Schneider D, et al. (2011) Negative epistasis between beneficial mutations in an evolving bacterial population. *Science* 332:1193–1196. doi: 10.1126/science.1203801
- Kvitek DJ, Sherlock G (2011) Reciprocal sign epistasis between frequently experimentally evolved adaptive mutations causes a rugged fitness landscape. *PLoS Genet* 7:e1002056. doi: 10.1371/journal.pgen.1002056

Formatiert: Deutsch (Schweiz)

- Layzer D (1978) A macroscopic approach to population genetics. *Journal of Theoretical Biology* 73:769–788. doi: 10.1016/0022-5193(78)90136-4
- Lehman N, Joyce GF (1993) Evolution in vitro of an RNA enzyme with altered metal dependence. *Nature* 361:182–185. doi: 10.1038/361182a0
- Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer (Sunderland, Mass.)
- Pitt JN, Ferre-D'Amare AR (2010) Rapid Construction of Empirical RNA Fitness Landscapes. *Science* 330:376–379. doi: 10.1126/science.1192001
- Rutherford SL, Lindquist S (1998) Hsp90 as a capacitor for morphological evolution. *Nature* 396:336–342. doi: 10.1038/24550
- Scharloo W (1991) Canalization: Genetic and Developmental Aspects. *Annual Review of Ecology and Systematics* 22:65–93. doi: 10.1146/annurev.es.22.110191.000433
- Soskine M, Tawfik DS (2010) Mutational effects and the evolution of new protein functions. *Nat Rev Genet* 11:572–582. doi: 10.1038/nrg2808
- Tanner M, Cech T (1996) Activity and thermostability of the small self-splicing group I intron in the pre-tRNA(Ile) of the purple bacterium *Azoarcus*. *RNA* 2:74–83.
- Tokuriki N, Stricher F, Serrano L, Tawfik DS (2008) How Protein Stability and New Functions Trade Off. *PLoS Comput Biol* 4:e1000002. doi: 10.1371/journal.pcbi.1000002
- Traverse CC, Mayo-Smith LM, Poltak SR, Cooper VS (2013) Tangled bank of experimentally evolved *Burkholderia* biofilms reflects selection during chronic infections. *Proc Natl Acad Sci USA* 110:E250–259. doi: 10.1073/pnas.1207025110
- True HL, Lindquist SL (2000) A yeast prion provides a mechanism for genetic variation and phenotypic diversity. *Nature* 407:477–483. doi: 10.1038/35035005
- Turelli M, Barton NH (1990) Dynamics of polygenic characters under selection. *Theoretical Population Biology* 38:1–57. doi: 10.1016/0040-5809(90)90002-D
- De Visser JAGM, Hermisson J, Wagner GP, et al. (2003) Perspective: Evolution and detection of genetic robustness. *Evolution* 57:1959–1972.
- Waddington CH (1959) Canalization of development and genetic assimilation of acquired characters. *Nature* 183:1654–1655.

Formatiert: Französisch (Schweiz)

Wagner GP, Booth G, Bagheri-Chaichian H (1997) A Population Genetic Theory of Canalization. *Evolution* 51:329–347. doi: 10.2307/2411105

Wang X, Minasov G, Shoichet BK (2002) Evolution of an Antibiotic Resistance Enzyme Constrained by Stability and Activity Trade-offs. *Journal of Molecular Biology* 320:85–95. doi: 10.1016/S0022-2836(02)00400-X

Figure Legends

Fig. 1 The *Azoarcus* ribozyme and phenotype distributions. **a** The secondary structure of the *Azoarcus* Group I ribozyme is shown, including the P1 helix formed between the ribozyme and the oligonucleotide substrate. The cleavage site is indicated by an “x”, which is the site of the phosphorothioate bond in the substrate used for directional selection. **b** The chemistry catalyzed by the ribozyme. **c** The distribution of phenotypes in each population. From each population, 19 individual molecules were isolated and their activity was determined. The stabilizing selection (SS) population activities were determined with the *native* ssRNA substrate. The directional selection (DS) population were determined with the *new* phosphorothioate containing substrate. Activity was determined as the fraction of ribozyme reacted after 1h (see Methods). Each data point in the plot is the average of three replicates. The error bars show the mean and standard deviation of the sample

Formatiert: Schriftart: Fett

Formatiert: Schriftart: Fett

Formatiert: Schriftart: Kursiv

Formatiert: Schriftart: Fett

Formatiert: Schriftart: Kursiv

Formatiert: Schriftart: Fett

Formatiert: Schriftart: Fett

Fig. 2 The distribution of genotypes in sequence space. Sequences from the SS population (blue) and the DS population (red) were combined and then aligned to generate a large multiple sequence alignment. This alignment was decomposed using a principle component analysis. The three most informative components are plotted in a 3D-scatter plot

Formatiert: Schriftart: Fett

Fig. 3 Genotypic distance measurements of 1577 sequences from the SS population (blue) and 1830 sequences from the DS population (red). **a** Distance to wild-type. A pairwise alignment between each sequence and the wild-type sequence was performed. Distance was determined for the portion of each sequence (159 b.p.) that was inside the primer binding sites used to replicate molecules by PCR. The frequency (y-axis) with which each distance occurred is plotted. **b** Distributions of D_{within} measurements. Each sequence was aligned to every other non-identical sequence in the sample (All-against-all). Distance was determined for the 159 base pairs between the primer binding sites. The frequency (y-axis) with which each distance occurred in the sample is plotted. **c** Change in the variance of D_{within} measurements over the course of the selections. The variance of D_{within} was determined from sequence data from each round of stabilizing selection (blue) and directional selection (red). Each set of data was fit by linear regression (black lines). Forecasted variance values (grey) are for visual comparison, and were predicted based on the linear regression of the stabilizing selection data. Error bars represent the standard error of this regression

Formatiert: Schriftart: Fett

Formatiert: Schriftart: Fett

Formatiert: Schriftart: Fett

Formatiert: Schriftart: Fett